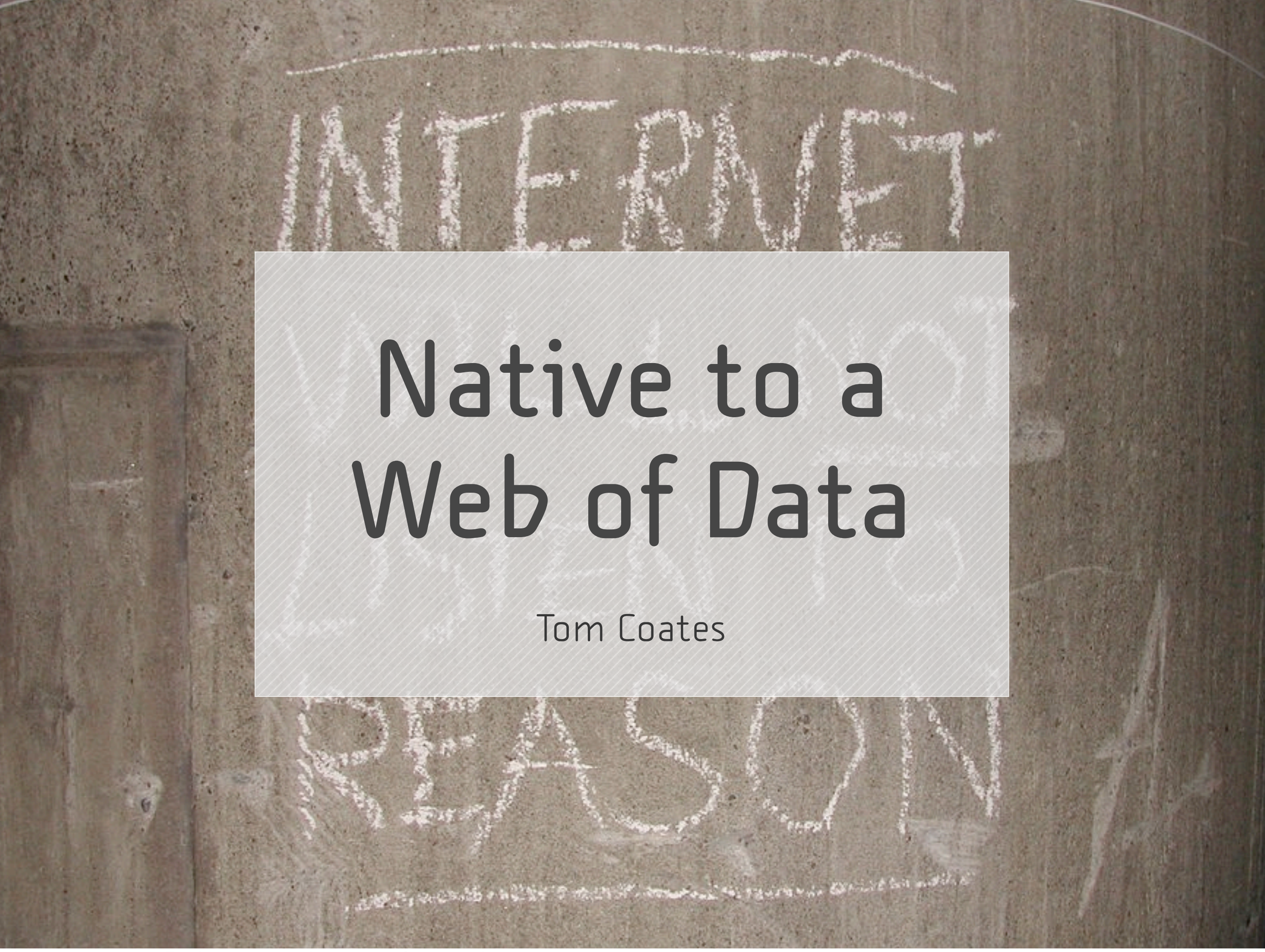


Native to a Web of Data

Tom Coates

http://www.flickr.com/photos/altemark/39593706/

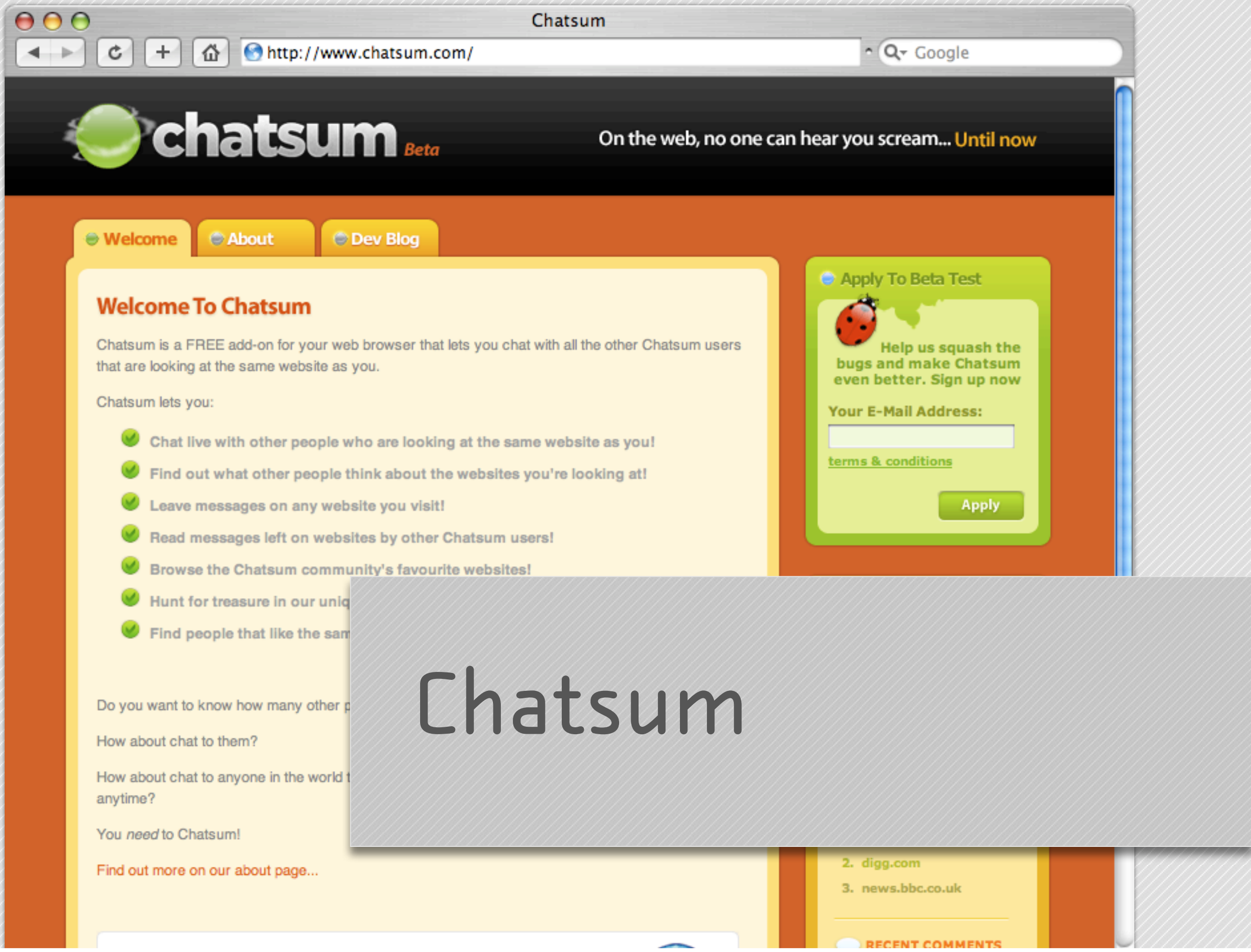
Hi, My name's Tom Coates. I've spent the last couple of years working for the BBC running a small R&D team in Radio and Music, and I'm currently doing something pretty similar for Yahoo, although I should state straight off that I've only been with them a couple of months and this talk is definitely my thoughts and not corporate policy.

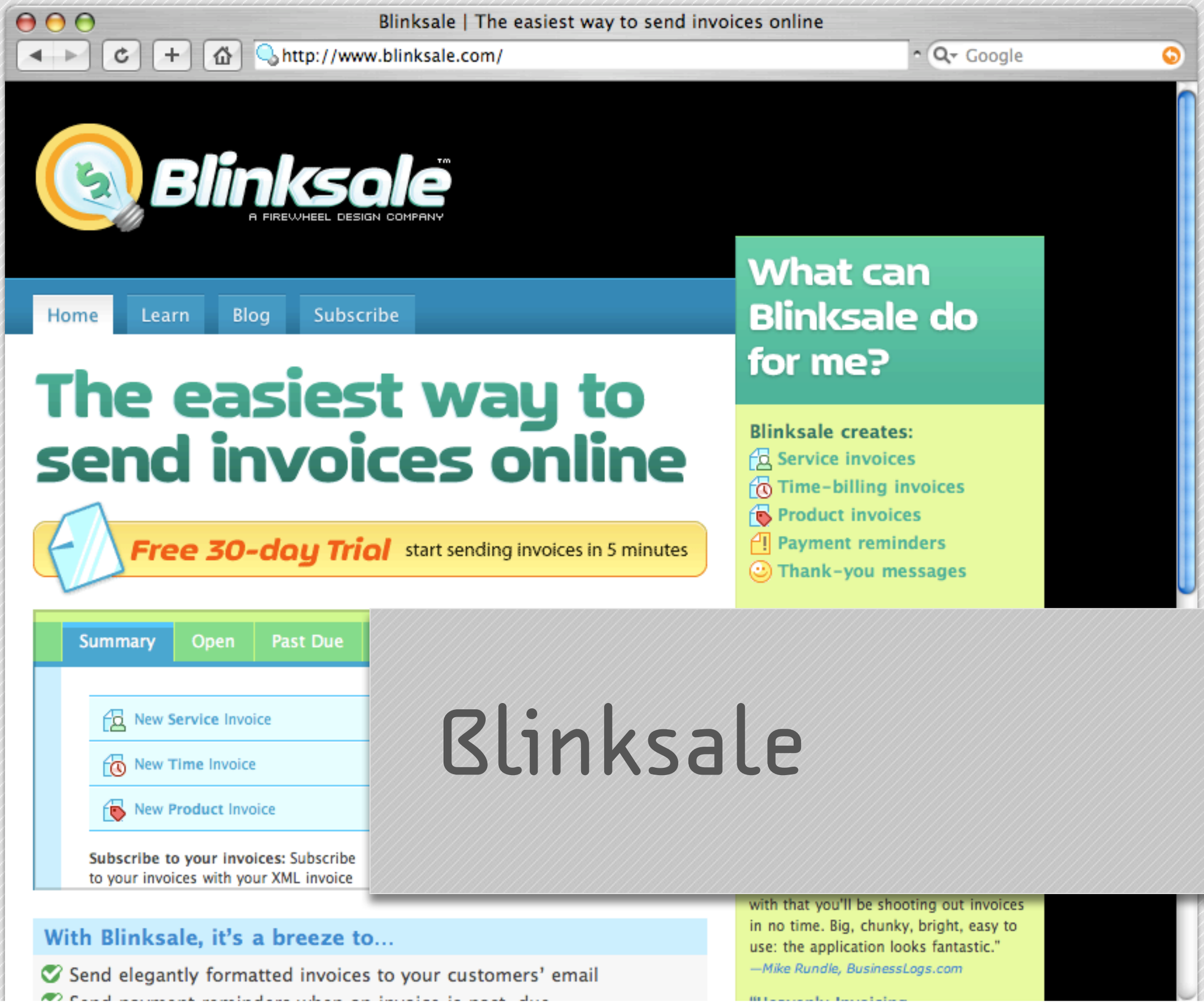
The only other part of my team in London at the moment is esteemed Python nerd Simon Willison, who is in the audience today. Ladies and Gentlemen, Simon Willison. Simon has been incredibly helpful in forcing me to get my head together for this presentation and I'd just like to start off by saying an enormous thank you to him, and to Matt Biddulph and Andy Budd who have been really cool.

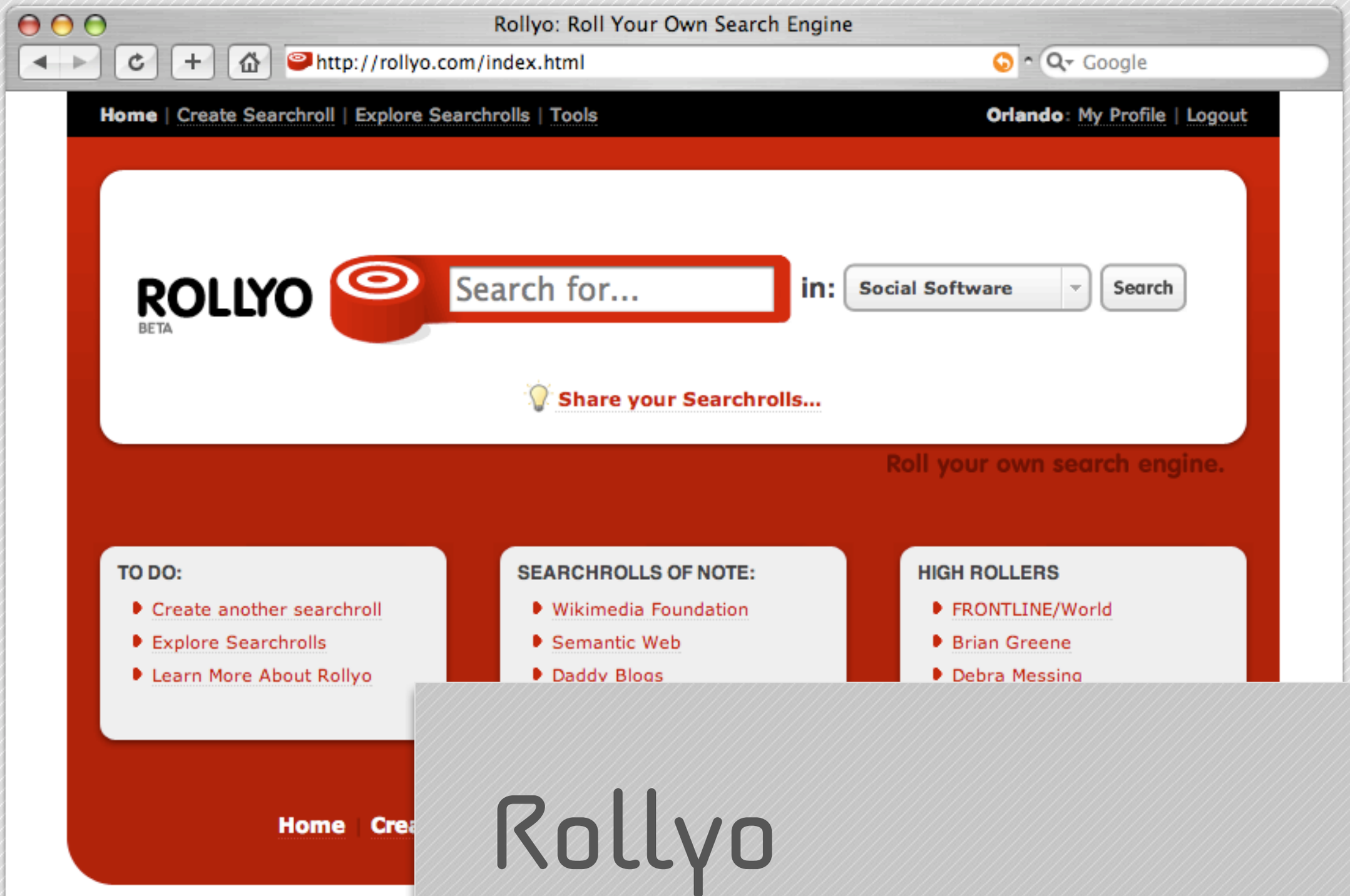
Anyway - as usual I've massively over-prepared, and I've got way way way too many slides and so I'm going to have to start. I'll be putting the whole thing online after the event so you can figure out what I was on about if I belt through it all a bit too fast...

Design and Web 2.0

Now, when you think about design and Web 2.0 probably the first thing you think of is bloody gradient fills and rounded corners.


Chatsum
http://www.chatsum.com/
Google
chatsum Beta
On the web, no one can hear you scream... Until now
Welcome
About
Dev Blog
Welcome To Chatsum
Chatsum is a FREE add-on for your web browser that lets you chat with all the other Chatsum users that are looking at the same website as you.
Chatsum lets you:
Chat live with other people who are looking at the same website as you!
Find out what other people think about the websites you're looking at!
Leave messages on any website you visit!
Read messages left on websites by other Chatsum users!
Browse the Chatsum community's favourite websites!
Hunt for treasure in our uniq
Find people that like the sam
Do you want to know how many other p
How about chat to them?
How about chat to anyone in the world t
anytime?
You need to Chatsum!
Find out more on our about page...
Apply To Beta Test
Help us squash the bugs and make Chatsum even better. Sign up now
Your E-Mail Address:
terms & conditions
Apply
2. digg.com
3. news.bbc.co.uk
RECENT COMMENTS
Chatsum


Blinksale | The easiest way to send invoices online
http://www.blinksale.com/
Google
Blinksale
A FIREWHEEL DESIGN COMPANY
Home
Learn
Blog
Subscribe
The easiest way to send invoices online
Free 30-day Trial start sending invoices in 5 minutes
What can Blinksale do for me?
Blinksale creates:
Service invoices
Time-billing invoices
Product invoices
Payment reminders
Thank-you messages
Summary
Open
Past Due
New Service Invoice
New Time Invoice
New Product Invoice
Subscribe to your invoices: Subscribe to your invoices with your XML invoice
With Blinksale, it's a breeze to...
Send elegantly formatted invoices to your customers' email
with that you'll be shooting out invoices in no time. Big, chunky, bright, easy to use: the application looks fantastic."
—Mike Rundle, BusinessLogs.com
Blinksale


Rollyo: Roll Your Own Search Engine
http://rollyo.com/index.html
Google
Home | Create Searchroll | Explore Searchrolls | Tools
Orlando: My Profile | Logout
ROLLYO
BETA
Search for...
in:
Social Software
Search
Share your Searchrolls...
Roll your own search engine.
TO DO:
Create another searchroll
Explore Searchrolls
Learn More About Rollyo
SEARCHROLLS OF NOTE:
Wikimedia Foundation
Semantic Web
Daddy Blogs
HIGH ROLLERS
FRONTLINE/World
Brian Greene
Debra Messing
Home | Crea

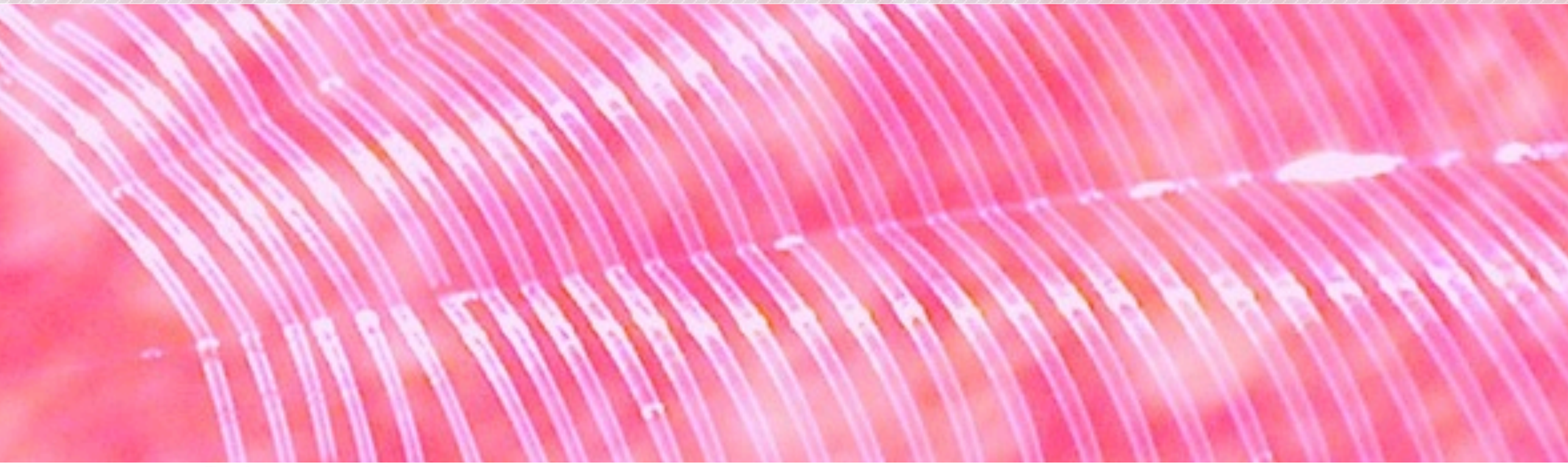
Rollyo

But I'm not going to be talking about rounded corners and gradient fills at all. I'm going to be talking about product design at a higher level – about what it means to build a product that fits and works and thrives in its environment. Because the web as an environment is starting to change quite dramatically, it's starting to become more than the sum of its parts.

The next 41 minutes

- What is the web changing into?
- What can you / should you build on it?
- Architectural principles for Web 2.0

A web of data

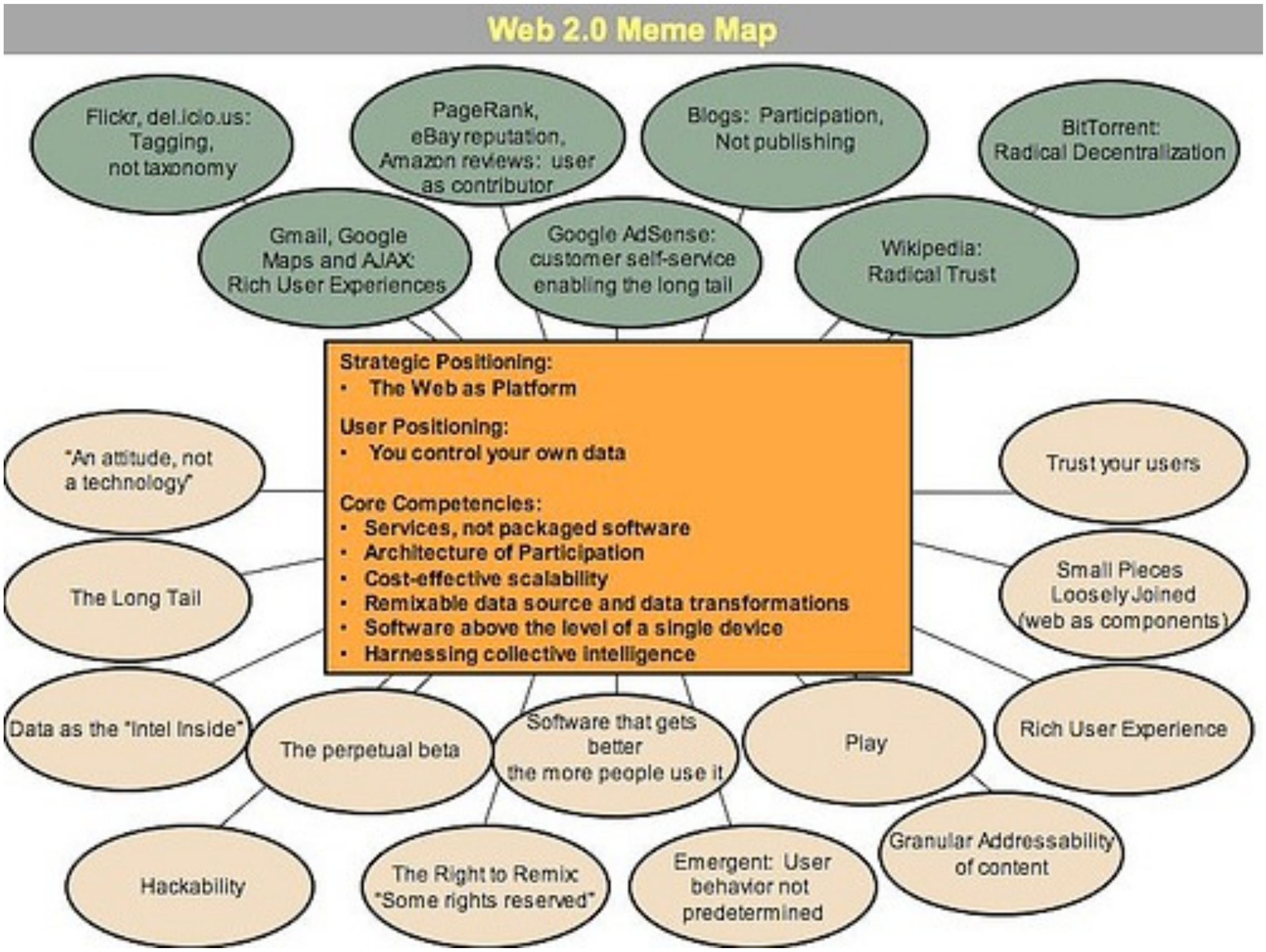


http://www.flickr.com/photos/777/70455377/ – Naomi Ibuki

Web 2.0?

- Buzzword
- Conference
- Marketing
- A new bubble
- All of these and less?

Well it's easy to be cynical about these buzzwords, but the truth is that we've been seeing a mass of new developments over the last couple of years. Before the bubble, the internet was about communication, then with the arrival of the big companies it became all about sales and publishing. But things have been changing again for a few years and we're now in a stage when people are trying to make sense of it. Web 2.0 as a phrase is just one of those attempts.



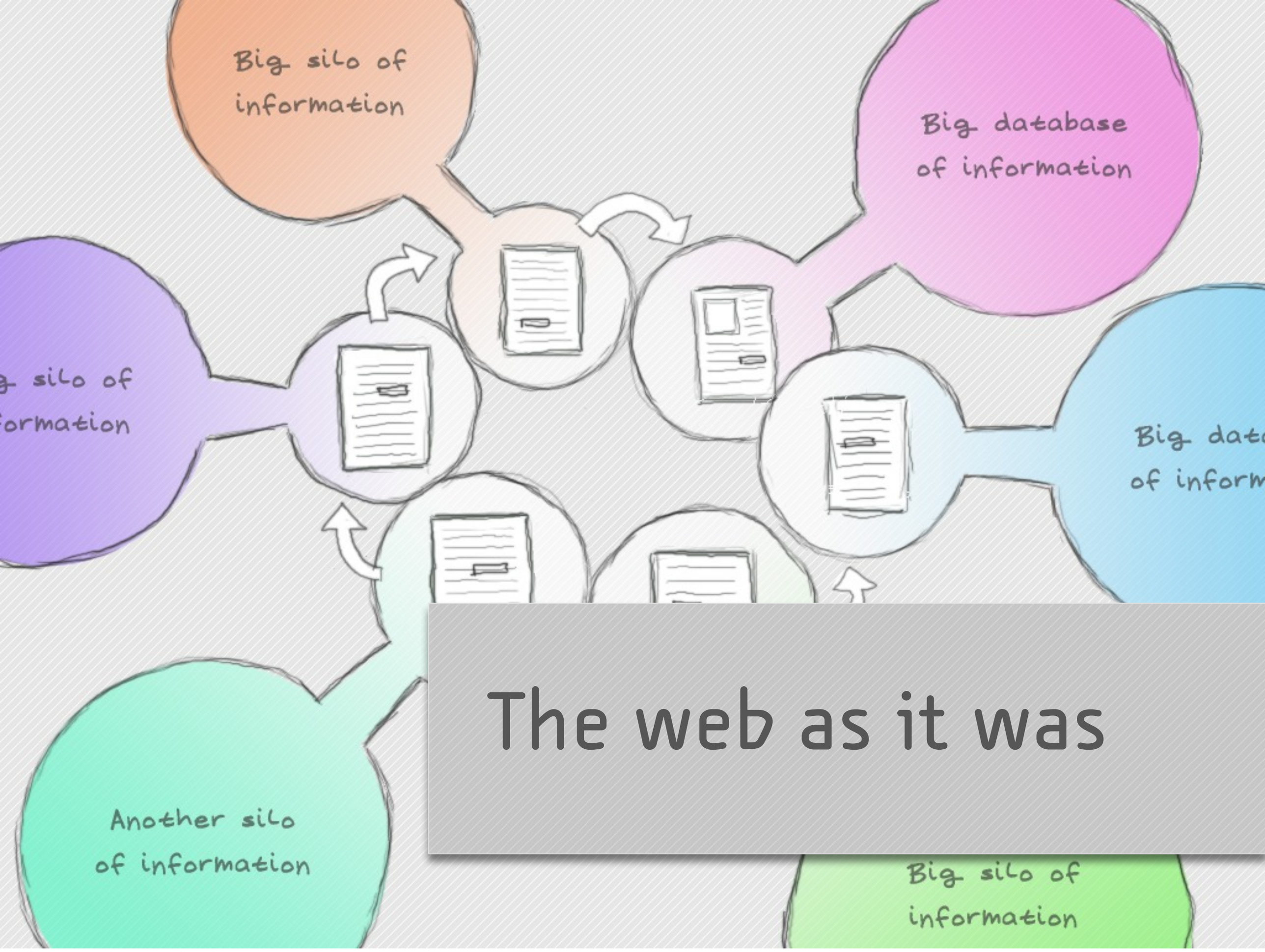
Probably the most famous attempt to make sense of all the trends that are happening at the moment is Tim O'Reilly's piece on "What is Web 2.0" based on a session he hosted at last year's FOO Camp. Tim and the O'Reilly group also coined the Web 2.0 term. I can really recommend the article - Web 2.0 as a term generates some scepticism from people, but the article itself is way better than most of the commentary that followed.

But I think I prefer this representation of Web 2.0, as made by Markus Angermeier at the end of last year. It's a kind of mind-mappish attempt to pull together some patterns, and I think gestures at how many things seem to be changing and pushing forward concurrently.

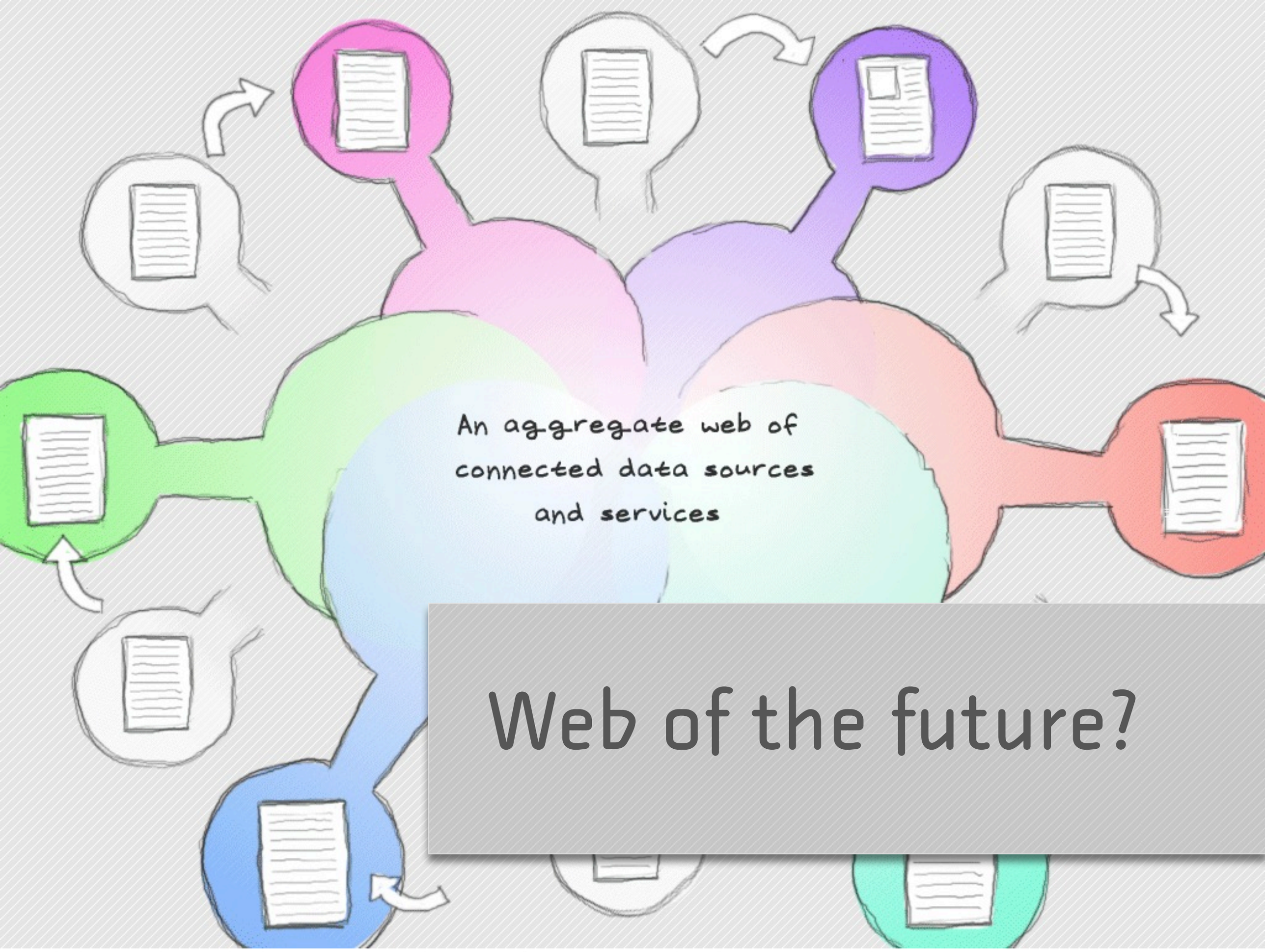
I'd argue though, that the term Web 2.0 isn't strong enough to hold together all these disparate streams. You've got a whole lot of activity going on here at different levels - architectural changes, cosmetic changes, environmental changes, technological, social and business changes - and I think it's over-ambitious to make it all representative of one underlying process.

So I'm going to personally abandon this term and concentrate on a subset where I think the web is experiencing some really dramatic changes, and I use the web here really carefully. Because you only have to look at Google and Yahoo and Flickr and del.icio.us and the various other web start-ups to see that the idea of a web of connected resources has developed pretty quickly over the last few years beyond the simple idea of the navigational hyperlink between published pages and - through APIs, web services and RSS - right down into the heart of the data itself.

We've got a new web manifesting here, a web of connected data. So for example, if this is the web that was....



This is the web that was...



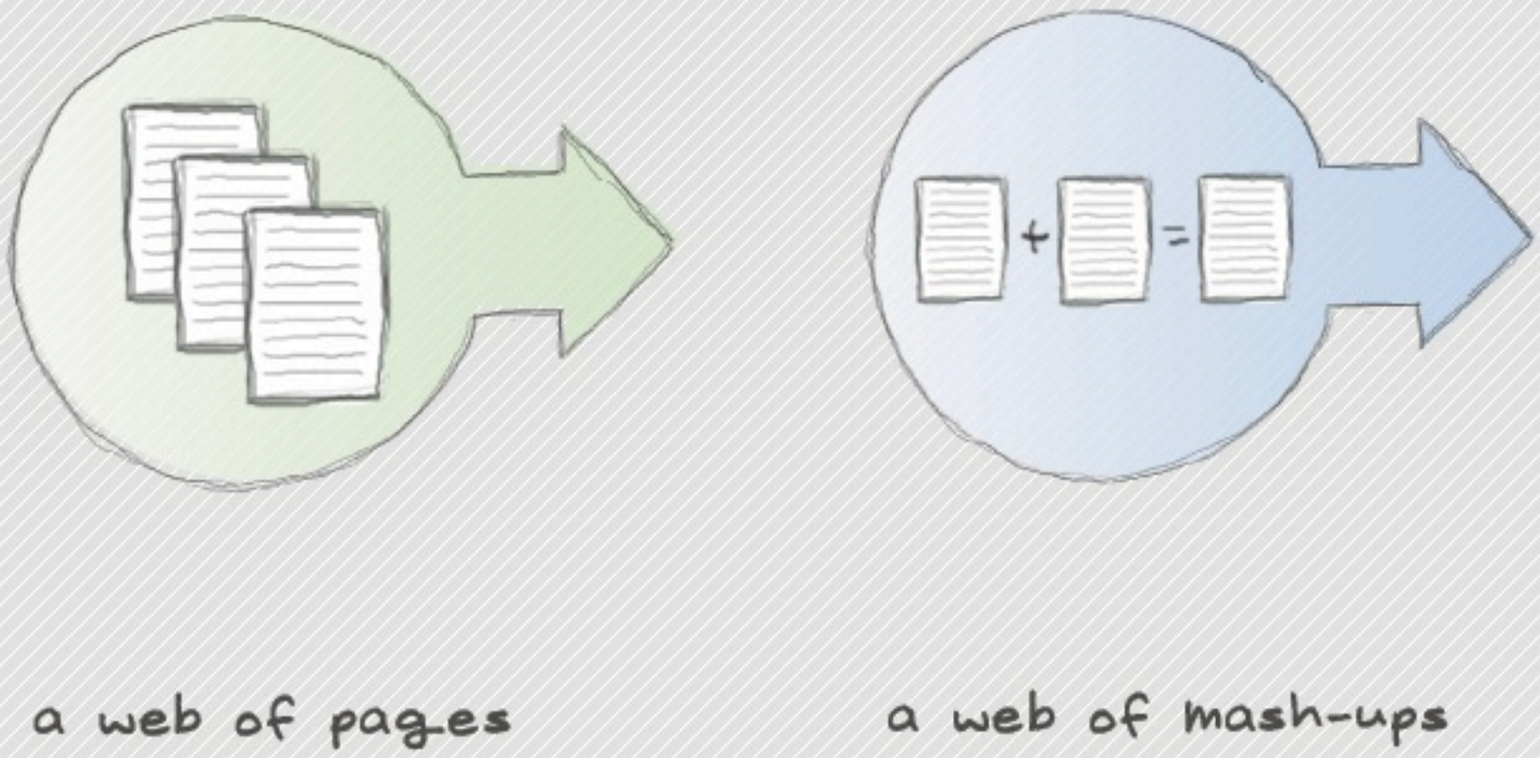
And this is the web that seems to be manifesting. A web of connected data resources, each of which connected to the others around them, able to create more by their combination than they could apart.

This is the web I think we're moving towards and it's very different from the web we have now. Unlike today, where we have pretty much a web of pages, we're on our way towards something like a web of data...

A web of data sources, services for exploring and manipulating data, and ways that users can connect them together

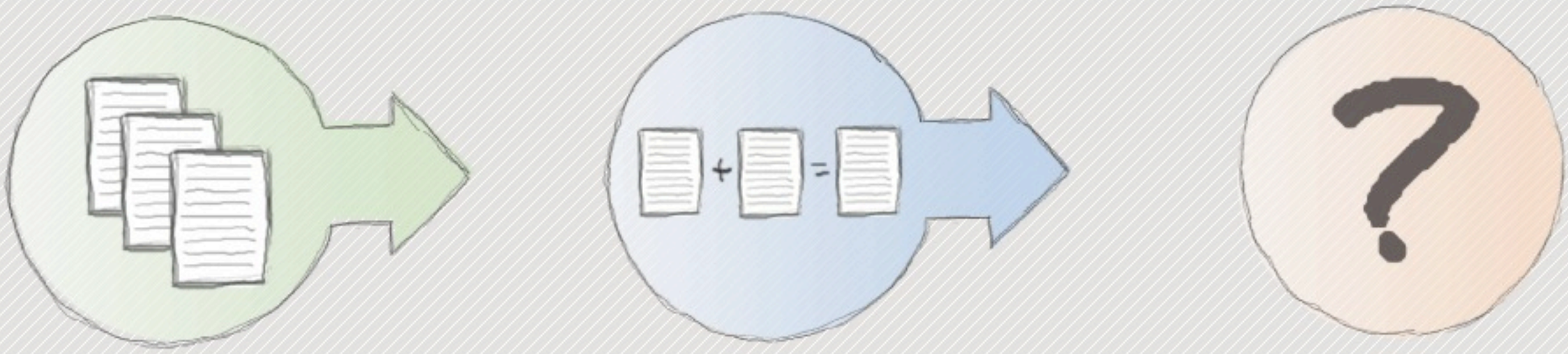
So this is what I'm talking about – a product of the gradual exposure of databases and APIs.

Very important to run through this slide, explaining what I mean by data (most importantly). The other two should become more clear as we work our way through it...



Mash-ups are kind of our pilot fish for whatever the web is becoming. In and of themselves they're fairly interesting, but if you think of them as the beginning of a trend towards interconnected data and reuse they're pretty fascinating...

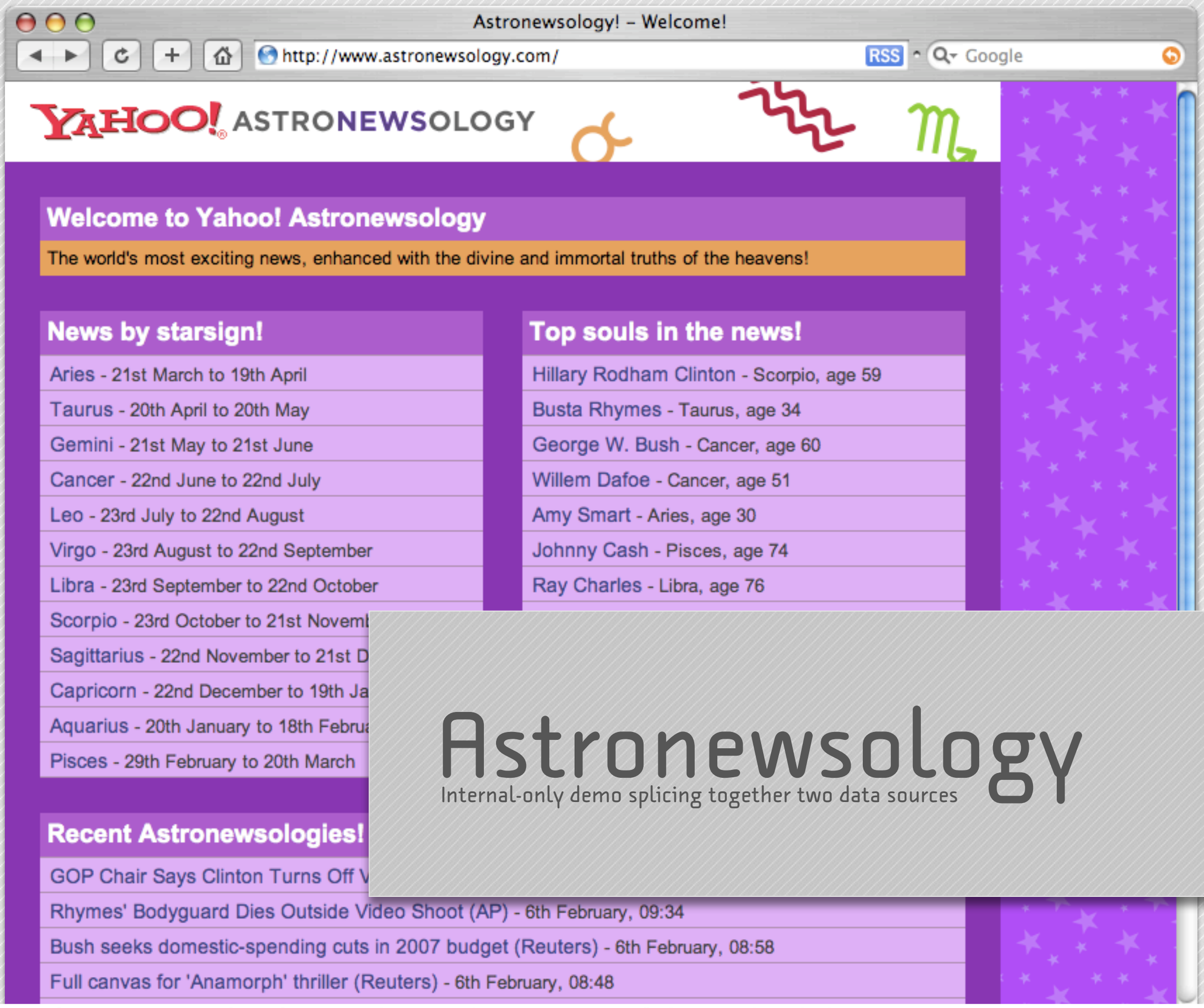
{Should probably say something here about creative commons and reuse, however briefly}



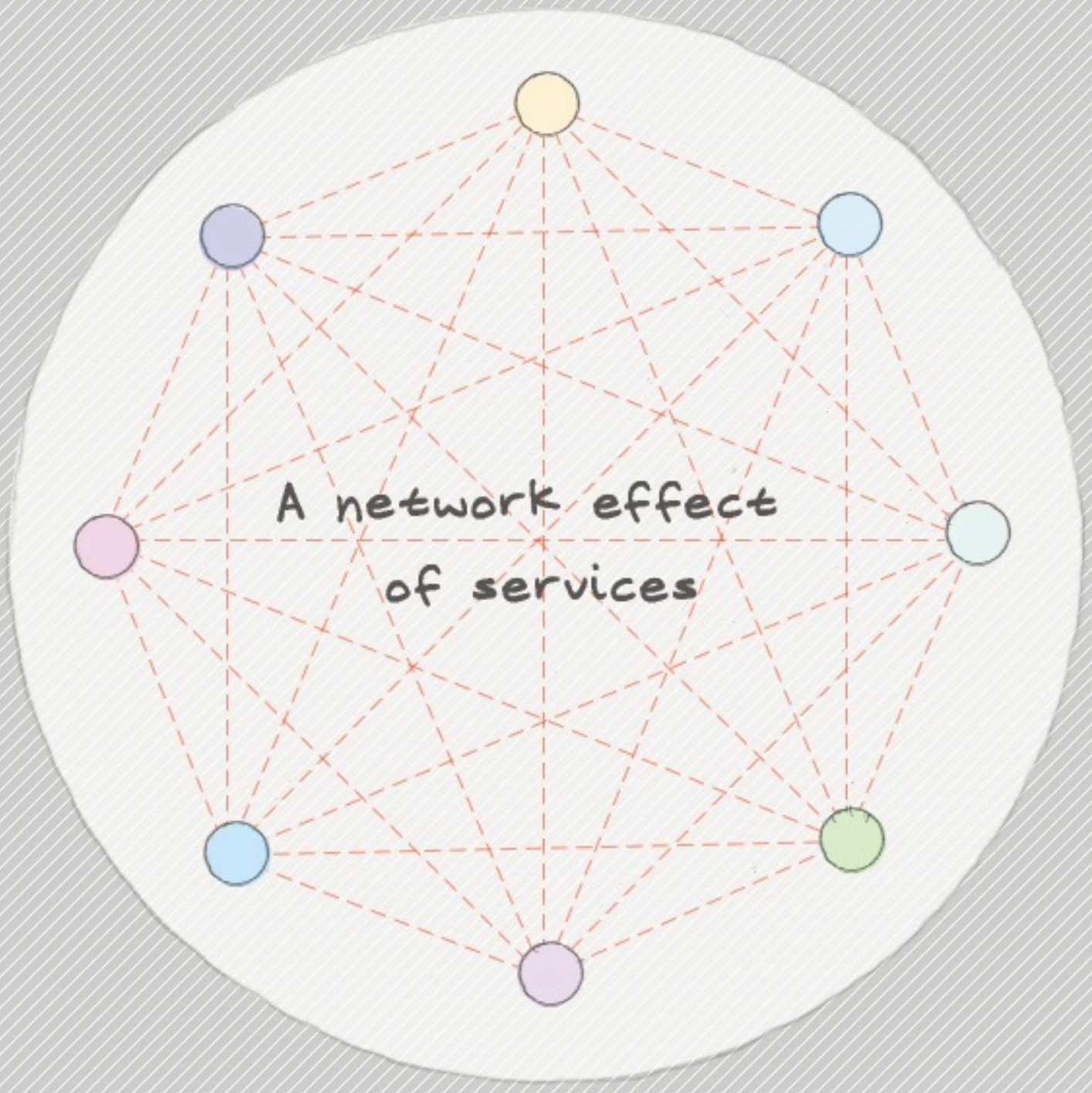
a web of pages

a web of mash-ups

a web of data



Astronewsology (internal only) combines three types of data – birthdays, star-signs and news stories


A network effect
of services

A Network Effect of services

- Every new service can build on top of every other existing service - the web becomes a true platform
- Every service and piece of data that's added to the web makes every other service potentially more powerful

Consequences

- Massive creative possibilities
- Accelerating innovation
- Increasingly competitive services
- Increasingly componentised services
- Increasingly specialised services

There is money to be made

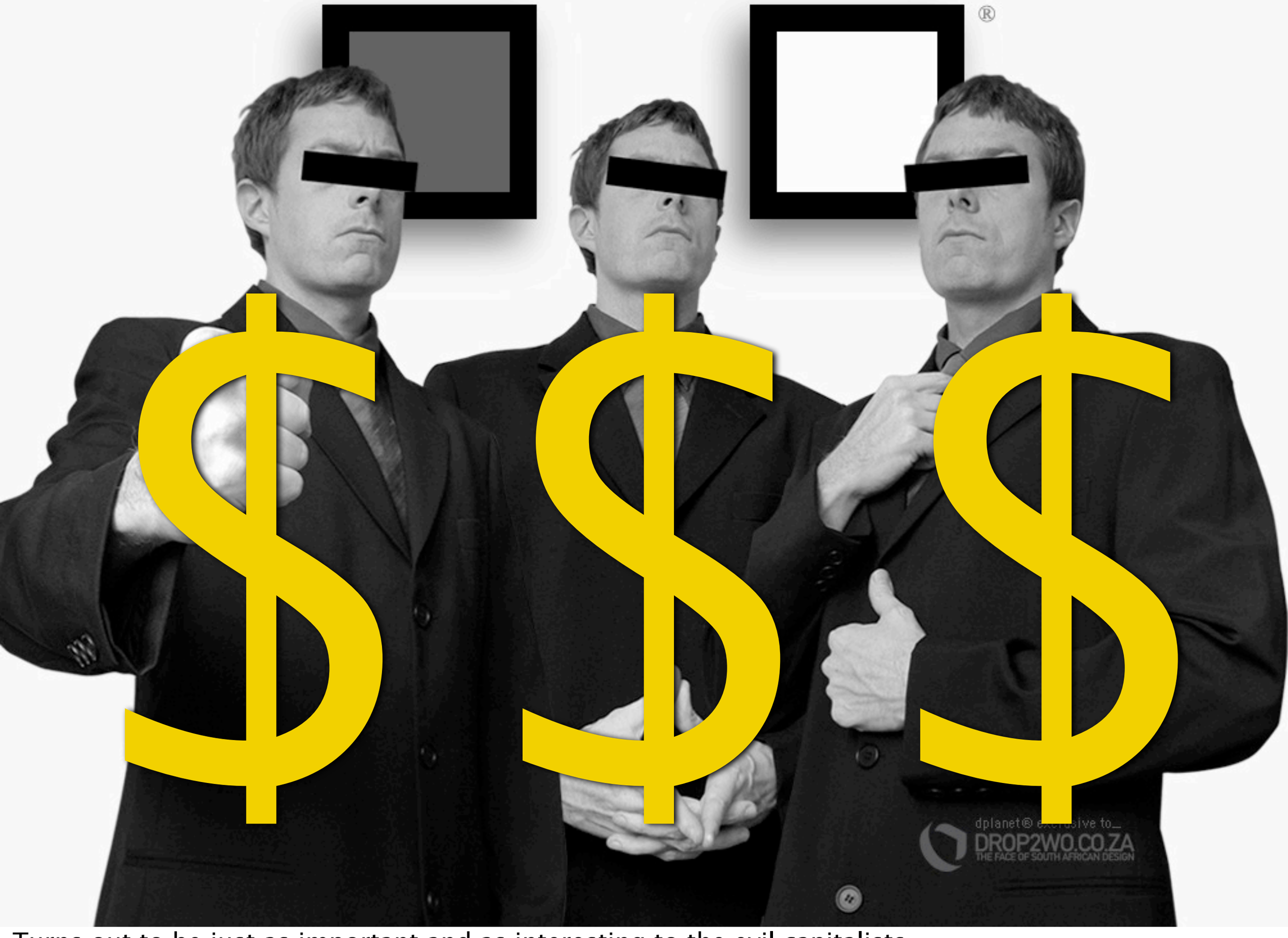
- Use APIs to drive people to your stuff
- Make your service more attractive and useful with less central development
- Use syndicated content as a platform
- Turn your API into a pay-for service

{outline above and then}

Basically, the same thing that we have traditionally associated with Creative Commons and radical idealists and playful people...

http://www.flickr.com/photos/nodditect/81114201/

Basically leftie liberals and hippies like, well, me...

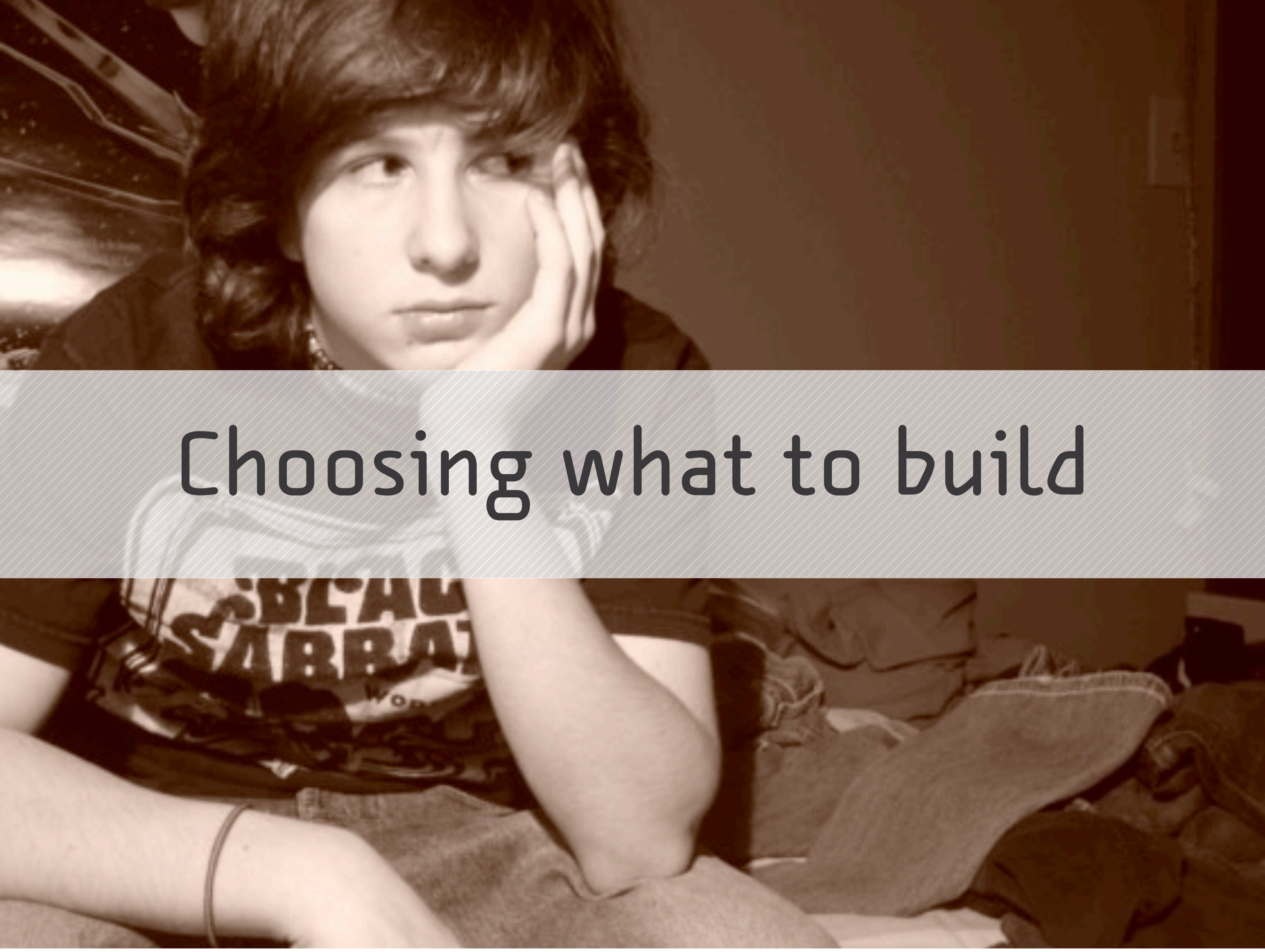


Turns out to be just as important and as interesting to the evil capitalists.

Which makes me even more certain that things are going to move in this direction. There's just a new playing field which we're all going to have to play on - for good or ill.

Because if you're not benefitting from the accelerating change, network effects and added value of being part of the web of data, then most likely you're going to be in backwater...

http://www.flickr.com/photos/dplanet/81933878/



Choosing what to build

So my argument is that if you're part of this ecosystem you will be pulled along and caught up in a web of accelerating value and reuse, whereas if you are not you'll be stuck in a disconnected backwater.

But what kinds of products work well in this space? How do you decide what to build?

http://www.flickr.com/photos/merfam/71578647/ - merfam

What can I build that will make the **whole Web** better?

How can I add value to the **Aggregate Web**?

A web of data sources, services for exploring and manipulating data, and ways that users can connect them together

So the first question is - can you find way to add data to the aggregate web? Can you create or open up a database of information that already exists, become a definitive home for a particular kind of data on the web? Can you own a kind of data that people want or are prepared to pay for? Can you work with the wider web to help your users create data, to publish or annotate or enhance some things that are already there? Or help them organise a part of their lives, help them turn their own information into data, and share and use it in more powerful ways?

Tim O'Reilly, "What is Web 2.0?"

"The race is on to own certain classes of core data: location, identity, calendaring of public events, product identifiers and namespaces."

"In many cases ... there may be an opportunity for an Intel Inside style play, with a single source for the data. In others, the winner will be the company that first reaches critical mass via user aggregation, and turns that aggregated data into a system service."

To be more cynical and businesslike, one shift in this ecosystem is going to be towards people trying to control and own certain key types of data, or to become synonymous with it. Tim O'Reilly says it best in this particularly blunt couple of quotes from the What is Web 2.0 piece...

A web of data sources, services for exploring and manipulating data, and ways that users can connect them together

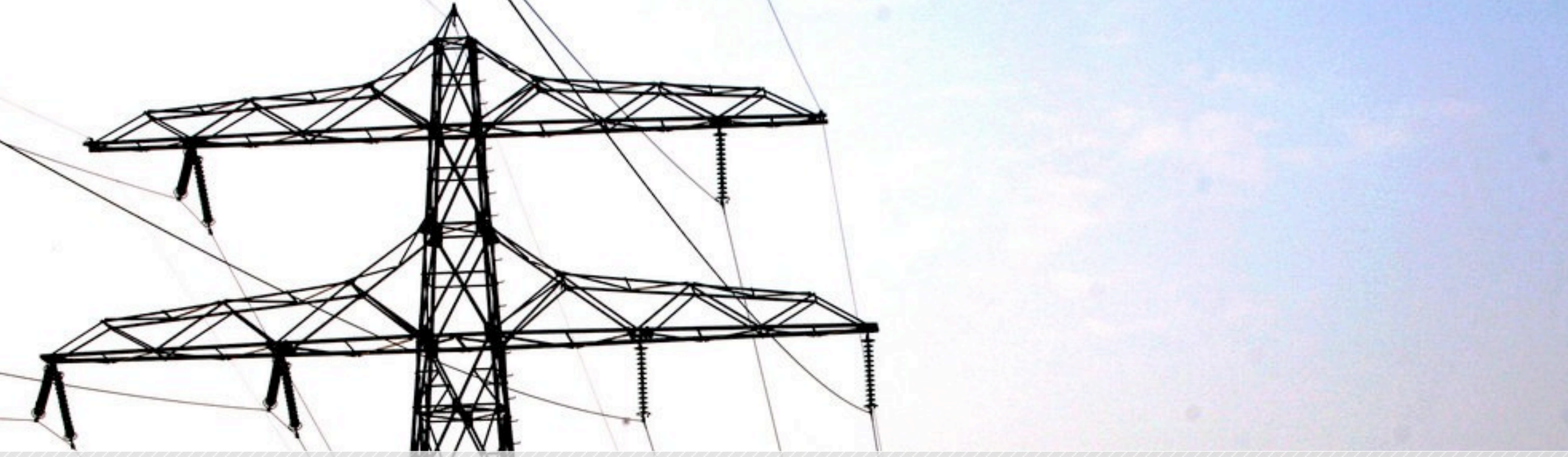
The next way you can add value is by making a service that helps people explore, use or manipulate data in some way. The arrival of weblogs and the popularisation of RSS amounted to pretty much the first improvement to structured data publishing on the popular internet for a long time and loads of people have built on top of the data they've created. Similarly Amazon web services and the Flickr APIs have created an incredibly fertile space both for individuals to play creatively and – increasingly – to build businesses on top of other people's data stores.

Examples: Aggregators? Flickr navigation interfaces?

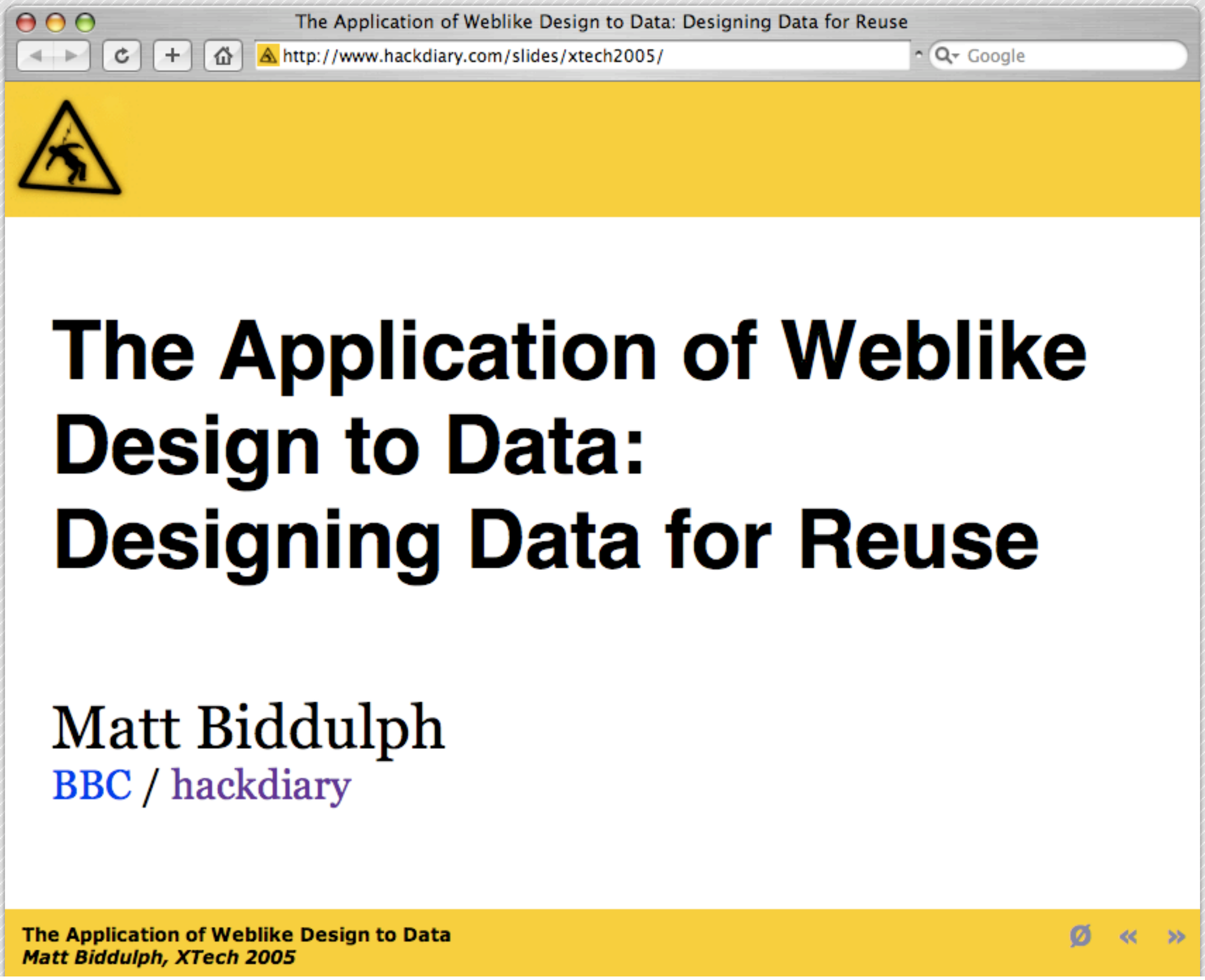
A web of data sources, services for exploring and manipulating data, and ways that users can connect them together

At it's most basic - can you move from people or organisations from capturing and organising their data into making it a more embedded part of this data ecosystem. Can you help them syndicate, help them cross publish - ideally without any extra work - and show them the benefits of being able to connect one sort of data with another?

Examples: Feedburner? All consuming?



Architectural principles


The Application of Weblike Design to Data: Designing Data for Reuse
http://www.hackdiary.com/slides/xtech2005/
Google
The Application of Weblike Design to Data: Designing Data for Reuse
Matt Biddulph
BBC / hackdiary
The Application of Weblike Design to Data
Matt Biddulph, XTech 2005

Some core components

- Data sources
- Standard ways of representing data
- Identifiers & URLs
- Mechanisms for distributing data
- Ways to interact with / enhance data
- Rights frameworks & financial

1

Look to add value to the Aggregate Web of data

* Mission: Try and build something that adds value to the aggregate web - improving a data source, finding a new way to connect disparate data sources, build a new interface for manipulating data.

2

Build for **normal users**, **developers** and **machines**

User interface changes in a web of data, because you have at least three types of user – normal humans who are looking to explore or use the information on your site, developers who are looking for the hooks that they can use to build upon your service, and software that's been trained to look for common features and standards directly at the data level. To be part of a web of data you need to build for all of them.

3

Start designing with **data**, not with **pages**

Always think about what you're making in terms of data/information and not pages. This sounds like it would scare off real people, but actually the opposite is the case. Having a clear understanding of the information that a page represents is a good thing for normal human users. In both jobs, what you're trying to do is turn information into navigable, explorable, reusable, connectable units.

You're looking for a best of both world's scenario, where you have data that is rich and consistent enough for machines to work with reliably organised in ways that make it explorable and comprehendable to humans.

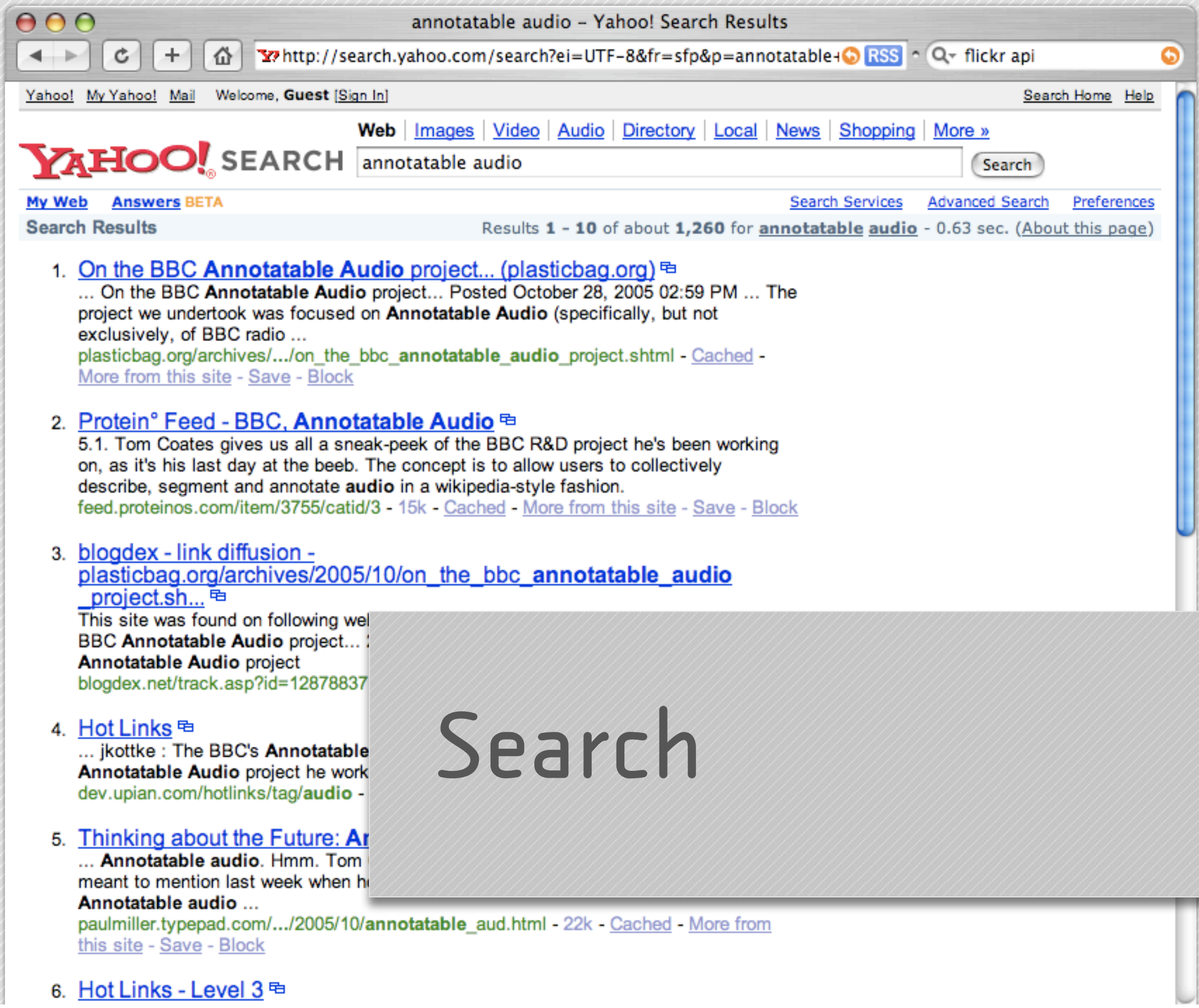
The process of product design starts with designing the data, and the data structures and relationships. If you don't capture information and relationships that will make it easy to navigate through your application or service then it will never work.

4

Identify your **first order** objects and make them **addressable**

Every page on your site will end up being an addressable view of your data. When I say addressable, I mean that you can point to it, or link to it, or send it to your friends or use it as a marker to stand in for the full contents of the page. So your first job is to understand what the core concepts are that you're going to be working with – whether it be people, addresses, events, photographs, television programmes or whatever – and to give each of these a unique and well structured URL.

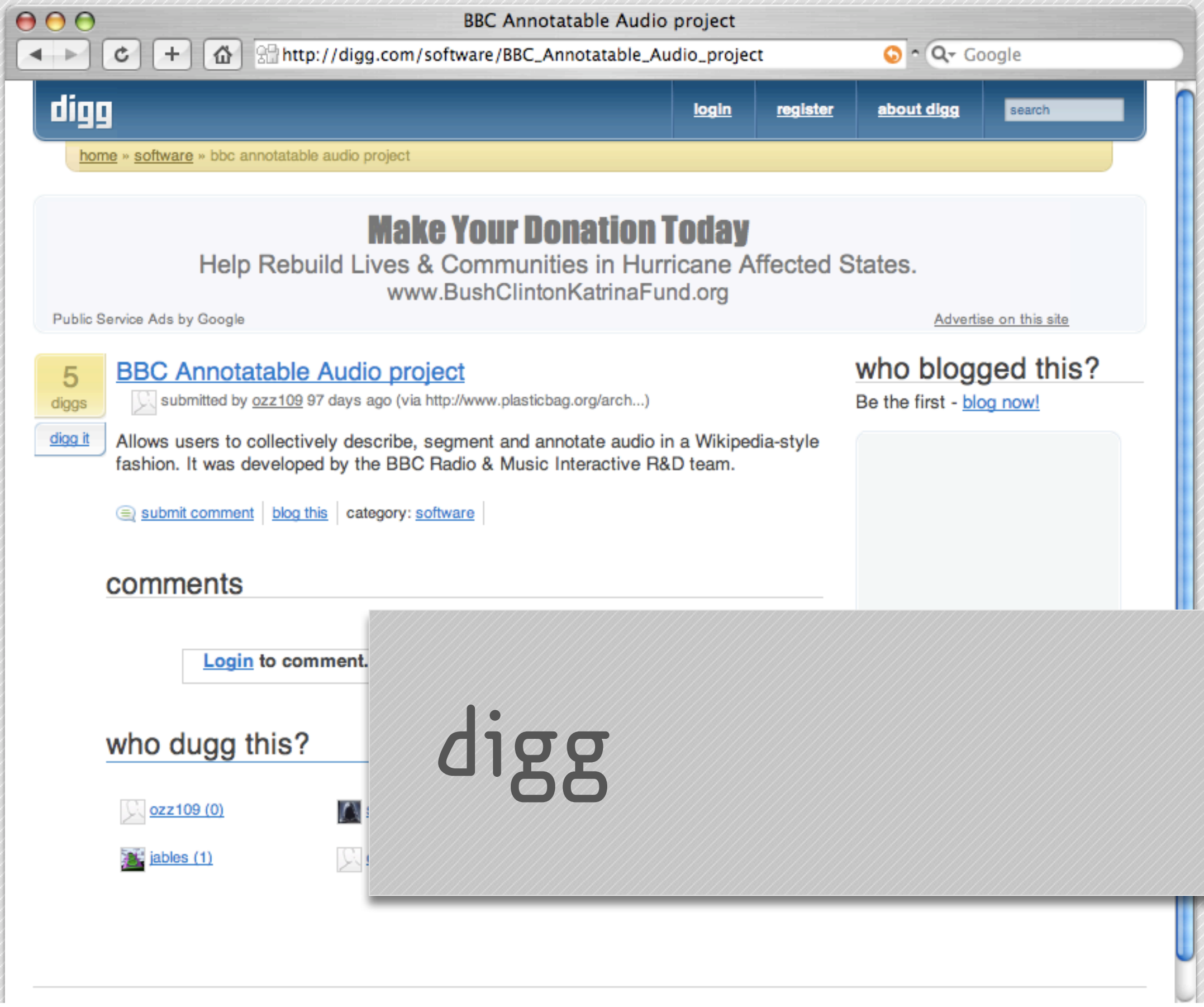
These will be your ‘destination’ pages.



Even if you do no more than that, you're already playing well in the aggregate web of data – search engines and aggregators like...

Technorati Search for: www.plasticbag.org/archives/2005/10/on_the_bbc_annotatable_audio_project.shtml
http://www.technorati.com/search/www.plasticbag.org%2Farchives%2F2
flickr api
Currently tracking 27.3 million sites and 2 billion links.
Member Sign In
Sign Up
Help
About
Technorati
Search
Tags
Blog Finder
Explore
www.plasticbag.org/archives/2005/10/on_the_bbc_annotatable_audic
Search
Options
What's this? This is a live search for posts that contain a link to www.plasticbag.org/arch..., sorted newest to oldest.
URL Search Help »
109 sites link to:
Add to Watchlist
View in Mini
www.plasticbag.org/arch...
Welcome back!
Do you blog? Claim it to get in here!
Want more? Become a member to save searches in a Watchlist.
3 posts in the last 4 days sorted by recency (sort by authority).
BBC Annotatable Audio Project
By Andy Polaine in Playpen 2 days ago
check this out
Fading Waypoints
In Fading Waypoints 4 days ago
that the interface threshold to create a tag is so low that people
ICANN ACCREDITED
Go Daddy.com
#1 OVERALL REGISTRAR
www.godaddy.com
Advertise on Technorati
US mourns the passing of Civil Rights leader Coretta Scott King. »
Today's Most Popular
News
BBC NEWS | South Asia | Fo...
CNN.com - Cartoon protests...
BBC NEWS | World | South A...
Movies
Brokeback Mountain
Munich (2005)
Walk the Line (2005)
Books
The Da Vinci Code
Harry Potter and the...
Memoirs of a Geisha
Technorati

technorati...


BBC Annotatable Audio project
http://digg.com/software/BBC_Annotatable_Audio_project
Google
digg
login
register
about digg
search
home » software » bbc annotatable audio project
Make Your Donation Today
Help Rebuild Lives & Communities in Hurricane Affected States.
www.BushClintonKatrinaFund.org
Public Service Ads by Google
Advertise on this site
5
diggs
digg it
BBC Annotatable Audio project
submitted by ozz109 97 days ago (via http://www.plasticbag.org/arch...)
Allows users to collectively describe, segment and annotate audio in a Wikipedia-style fashion. It was developed by the BBC Radio & Music Interactive R&D team.
submit comment
blog this
category: software
who blogged this?
Be the first - blog now!
comments
Login to comment.
who dugg this?
ozz109 (0)
jables (1)
digg

digg...

and del.icio.us can already start usefully aggregating information about each addressable component of your site, based on how people link to and reference them.

5

Use **readable**, **reliable** and **hackable** URLs

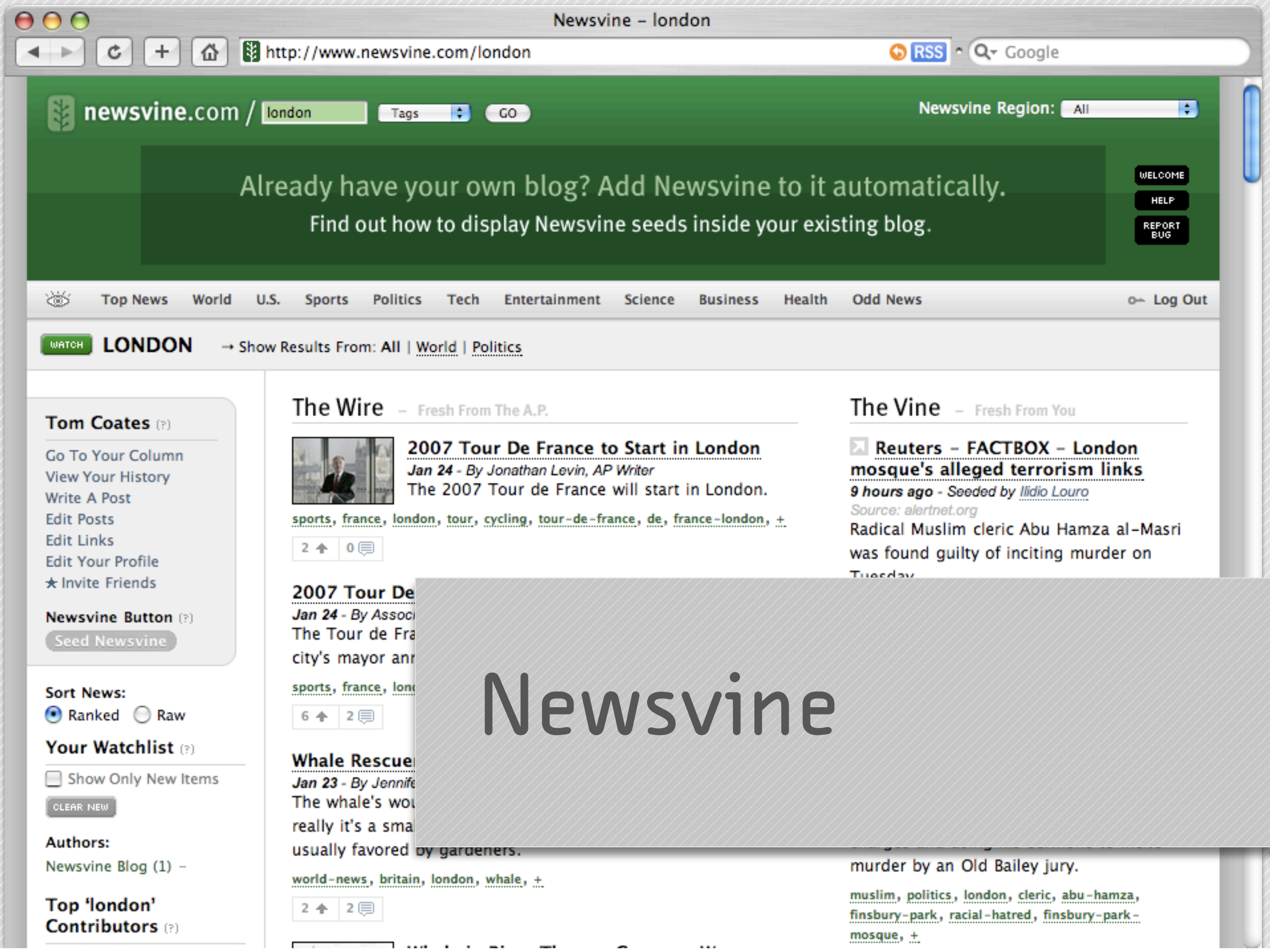
Good URLs should:

- be permanent references to resources
- have a 1-to-1 correlation with concepts
- use directories to represent hierarchy
- not reflect the underlying technology

Good URLs should:

- reflect the structure of the data
- be predictable / guessable / hackable
- be as human readable as possible
- be - or expose - identifiers

Some URL schemes are so elegant and powerful that they really offer themselves up as a major interface to the site itself – even to the extent that they have to be pulled into the page as a design element.



This kind of approach started on del.icio.us, but I thought it was really interesting that it found itself over on newsvine.

I'm not sure what I think of this approach, but it certainly shows you how powerful the URL can be in terms of supplementing or extending a site's navigation

... in addition creating a easy to automate way for a piece of software to connect and explore a site.

Good URLs are beautiful and a mark of design quality

Correlate with external identifier schemes (or coin a new standard)

But URLs aren't the only kind of addressability that you might need, because not all concepts are created natively in one place on the internet. A weblog post can be identified uniquely by it's URL, as can a photo on Flickr because they're not only the representation of that concept, they are the thing itself.

But what about films, tv shows, books, people, events? All of these things exist independently of the internet and are likely to have multiple and potentially competing representations online. You need a new concept to link all those representations together - to connect up the data produced in different places - and that's the idea of a unique identifier that represents that concept.

So, if you're working with types of data that already have a canonical or recognised authoritative representation or identifier exposed on the web, then build ways to correlate your identifier with the definitive one. If there aren't definitive URLs or identifiers out there already - which in many cases is more likely - then you will derive huge benefits from defining them or competing with dominant players who have coined them already.

7

Build list views and batch manipulation interfaces

* We've got our core first-order objects, and we've made them addressable, with a unique web page representing each one, and we've correlated those concepts with identifiers on the wider web.
* Now we have to think about ways in which you'd navigate between them, and ways in which you can manipulate and fiddle with the data you've got at your disposal, which is when we get to list views and ways of manipulating data...

Destination Page

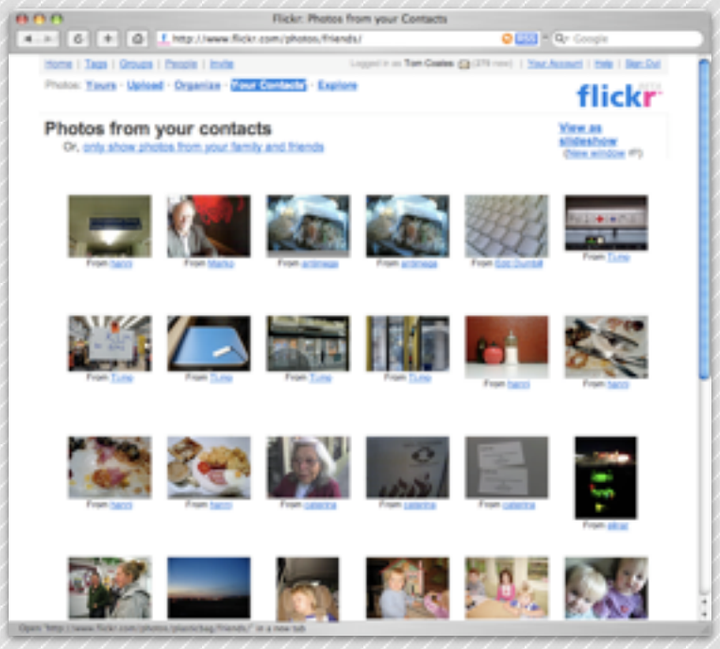
A core first-order concept and its subordinate information

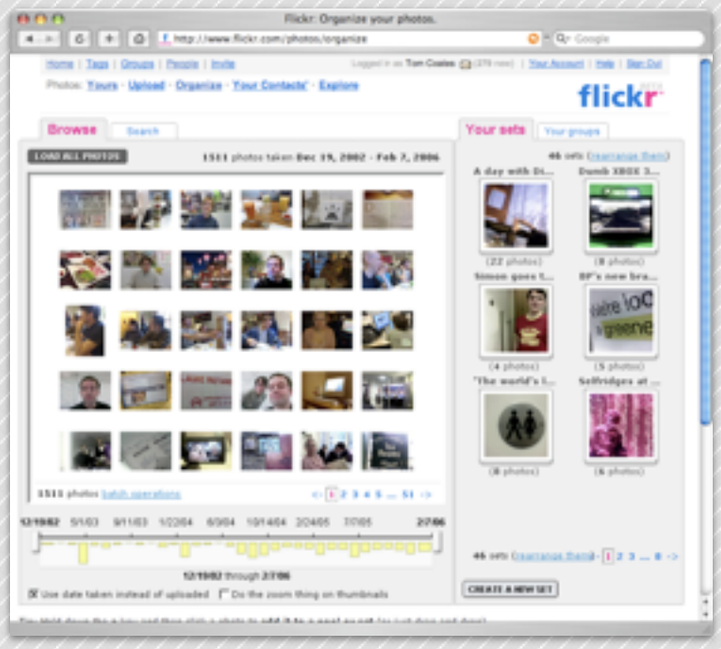
List-view page

A slice of your data used to navigate between first-order concepts

Manipulation Interface

Interface for the batch manipulation of first-order concepts





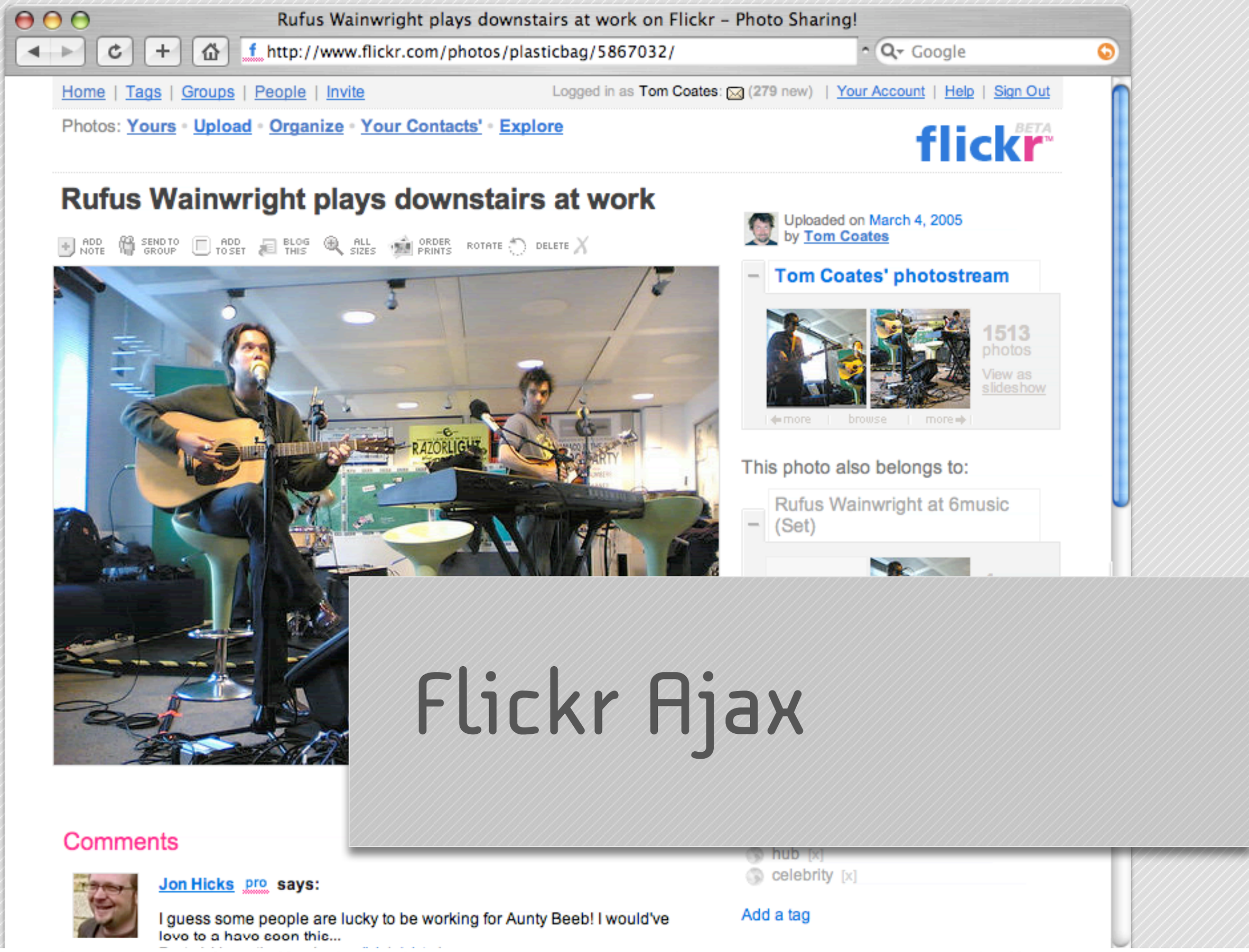
Three core types of page

There are fundamentally only really THREE CORE TYPES of pages that you need to build a web of data native service - or maybe even that's overstating it. It's possible you'll only need two. The core ones are:

* Pages representing your first order concepts - which is what we've talked about already, addressable concepts. On these pages, one of the bits of data that you'll want to have captured is explicit relationships to other first order concepts - ie. next in sequence and stuff like that.
* But the second type is about higher level views and lists of the first order objects - other ways of exploring that dataspace.
* The third form is really for convenience. If you're building a service where data manipulation is more core, and there's a lot of manipulation to do, then you might need a set of dedicated manipulation interfaces. Flickr has one in Organizr.

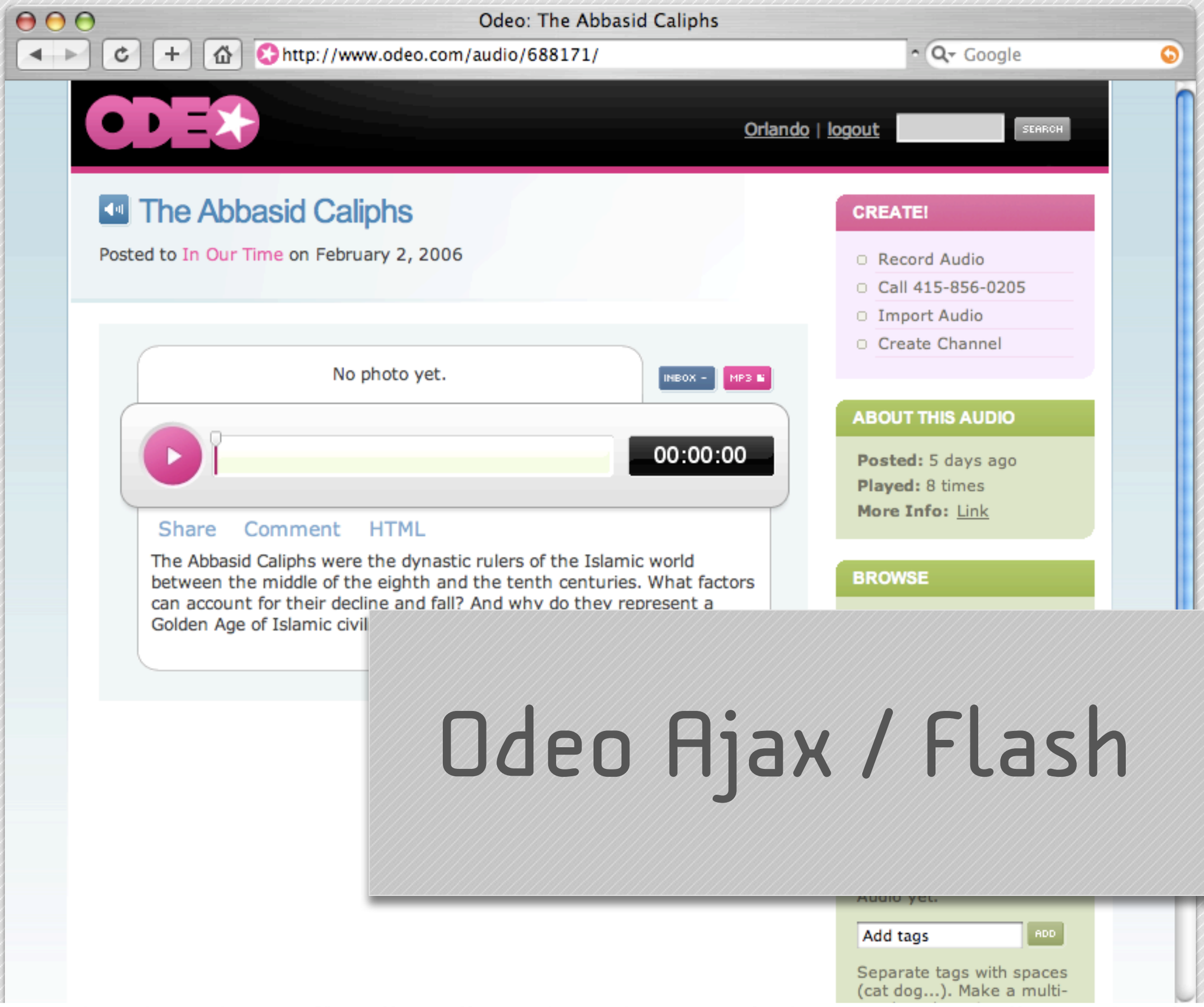
The concept of manipulation is really important and interesting one, because user manipulation of data is heavily constrained by the interface widgets at your disposal. Which is where new interface technologies like Flash and Ajax can come in.

The most important thing when using either technology is that you should absolutely *not* break the web. Each of your destination pages here should be addressable with a clear URL that represents a concept. Similarly each of your list view pages should have its own URL as well. Distinct things get distinct pages. Which means that if you're using Ajax or Flash on a page that's about a concept, you should only use it to help people manipulate or edit THAT CONCEPT. It's only in dedicated batch manipulation interfaces that you can go wild with this technology - because individual resources aren't necessarily supposed to be referenced while they're being fiddled with.



Flickr Ajax

Flickr does this extremely well - their destination pages each represent a photo and allow you to rotate the photo, add tags and annotate without refreshing the page. But the pages remain referenceable and part of the web. This kind of componentised use is a complete shift from the Flash / DHTML interfaces of yesterday and is all the better for it.



Odeo Ajax / Flash

Another site that does this stuff really well is Odeo. They use Flash components all over the place to allow in place previewing and listening to audio, but they never break

8

Create parallel data services using understood standards

Three views

Parallel data representations

Destination Page	List-view page	Manipulation Interface
A core first-order concept and its subordinate information	A slice of your data used to navigate between first-order concepts	Interface for the batch manipulation of first-order concepts

APIs (REST / SOAP ETC)

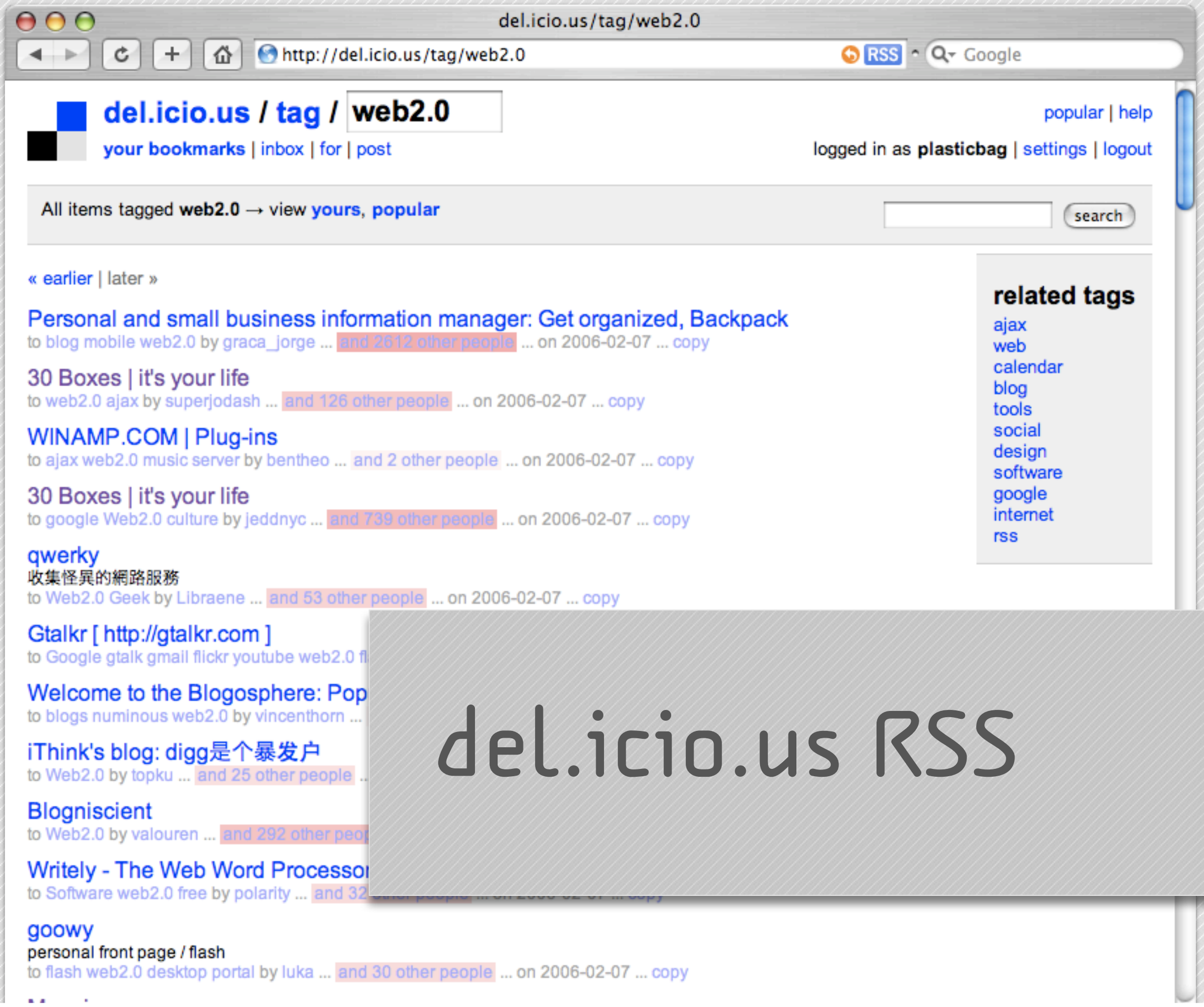
Microformats

Parallel XML

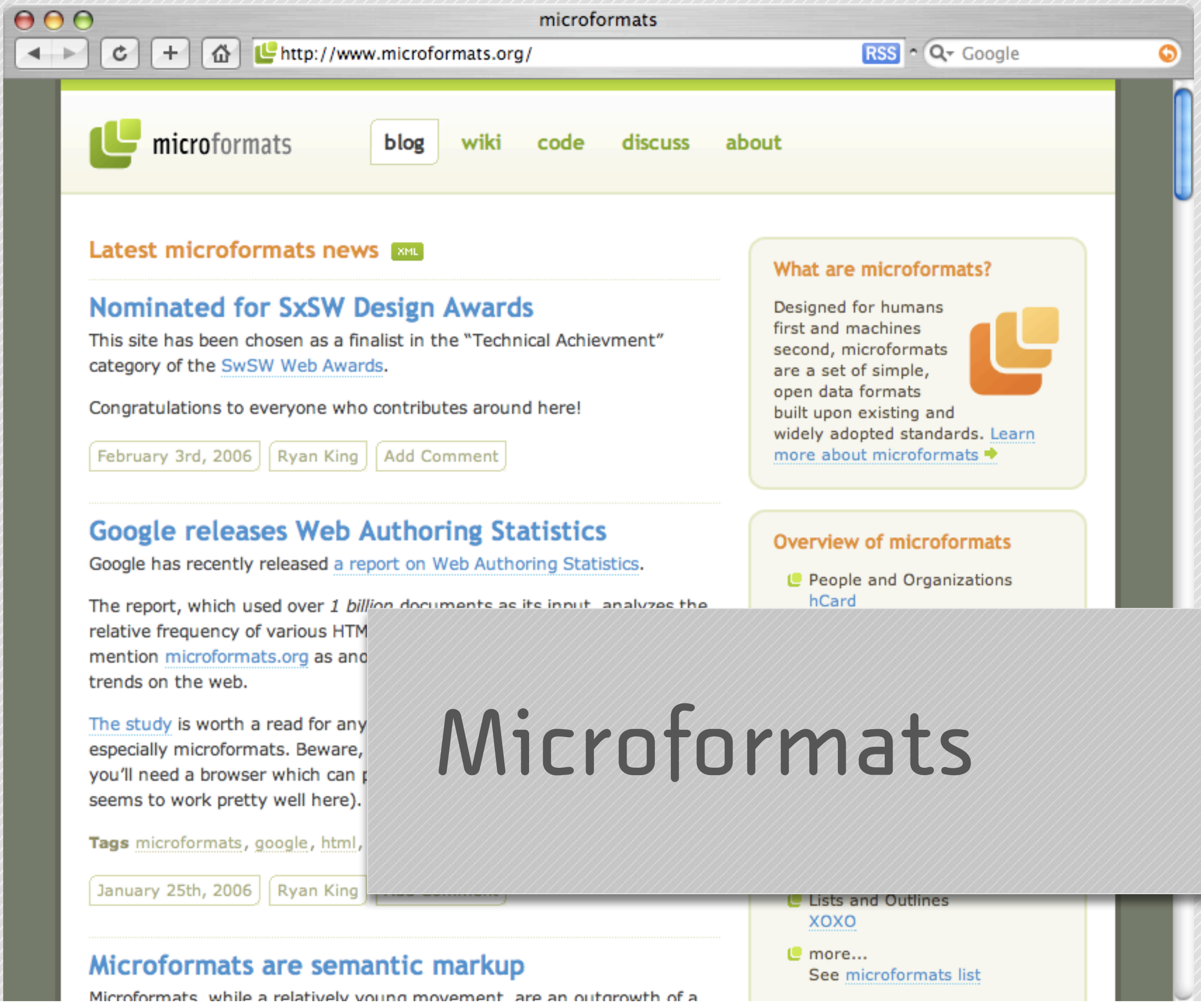
RSS

So you've got your three particular types of page

Ray Ozzie said in a leaked Microsoft Memo - RSS feeds (unix pipe of the internet?) / API.



del.icio.us RSS


microformats
http://www.microformats.org/
RSS
Google
microformats
blog
wiki
code
discuss
about
Latest microformats news
XML
Nominated for SxSW Design Awards
This site has been chosen as a finalist in the "Technical Achievment" category of the SwSW Web Awards.
Congratulations to everyone who contributes around here!
February 3rd, 2006
Ryan King
Add Comment
Google releases Web Authoring Statistics
Google has recently released a report on Web Authoring Statistics.
The report, which used over 1 billion documents as its input, analyzes the
relative frequency of various HTM
mention microformats.org as ano
trends on the web.
The study is worth a read for any
especially microformats. Beware,
you'll need a browser which can p
seems to work pretty well here).
Tags microformats, google, html,
January 25th, 2006
Ryan King
Microformats are semantic markup
Microformats, while a relatively young movement, are an outgrowth of a
What are microformats?
Designed for humans first and machines second, microformats are a set of simple, open data formats built upon existing and widely adopted standards. Learn more about microformats
Overview of microformats
People and Organizations
hCard
Lists and Outlines
XOXO
more...
See microformats list
Microformats

9

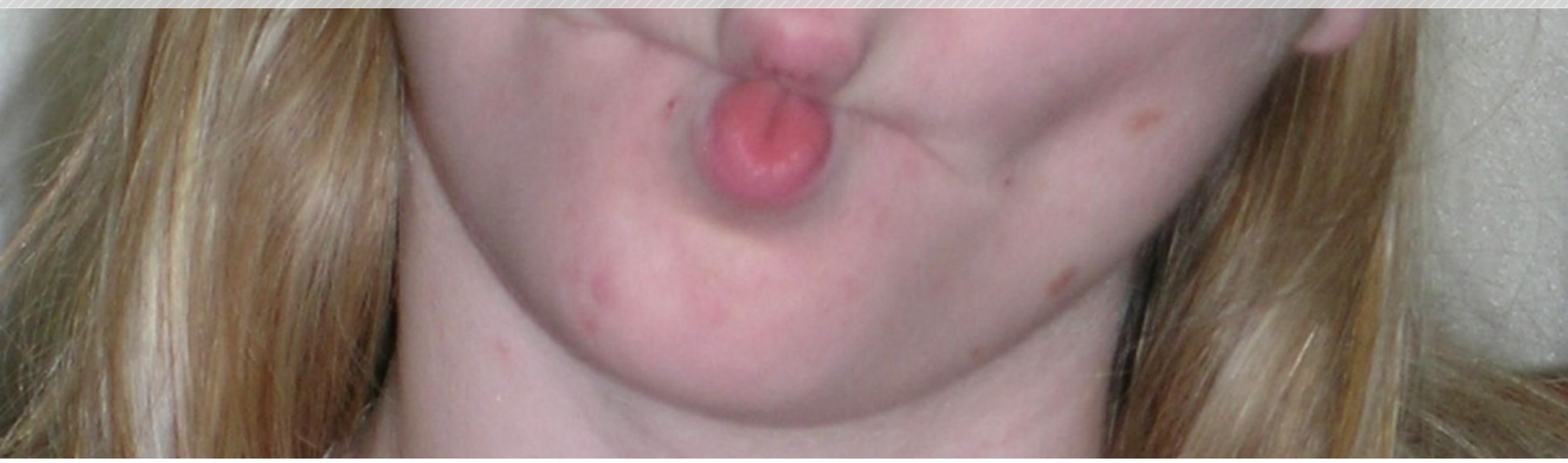
Make your data as discoverable as possible

Native to a Web of Data

1. Look to add value to the **Aggregate** Web of data
2. Build for **normal users**, **developers** and **machines**
3. Start designing with **data**, not with **pages**
4. Identify your **first order** objects and make them **addressable**
5. Use **readable**, **reliable** and **hackable** URLs
6. Correlate with external **identifier schemes**
7. Build **list views** and batch **manipulation interfaces**
8. Create **parallel data services** using **standards**
9. Make your data as **discoverable** as possible

§

If you've enjoyed this talk...



http://www.flickr.com/photos/angela7/76598852/

PLASTICBAG.ORG

A weblog by Tom Coates who thinks up neat stuff for Yahoo!
Concerning social software, mass-amateurisation, design, and future media
You can subscribe to an RSS feed, read the disclaimer or explore the archives

---

On the upcoming Carson Workshops summit...

Posted January 6, 2006 12:58 PM.

Quick announcement - I'm going to be talking at the upcoing Carson Workshops summit on The Future of Web Apps on the 8th of February in London. It's a one-day conference for developers and web application builders that's going to be focusing in on some of the technologies and ideas that are foundational to the web that is to come. It's got a pretty stellar group of people speaking - Joshua Schachter of del.icio.us will be talking about tags and how useful and important they are, David Heinemeier H
Osinga will be talking about G
and developing for Flickr, Ste
Shaun Inman will be talking ab
about Web 2.0 business mode

plasticbag.org

Yahoo! Developer Network Home – Welcome!

http://developer.yahoo.net/

Google

Developer Network Home Help

Site Search Search

YAHOO! DEVELOPER NETWORK

Yahoo! Developer Network Home - Welcome!

- Home
- **del.icio.us** *Save, tag, and share your favorite sites*
- **Finance** *Get stock quotes, company news*
- **Flickr™** *Share photos with your friends*
- **HotJobs** *Find a great job*
- **Maps** *Create rich embedded map applications and overlays*
- **Merchant Solutions** *Build, host, manage and promote your online store*
- **Music** *Customize the Yahoo! Music Engine*
- **RSS Feeds** *Integrate Yahoo! content with your site or app*
- **Search** *Integrate Web, image,*

Innovate.

You bring the skills. We bring the ingredients.

Welcome to the Yahoo! Developer Network. We help software developers integrate their Web sites and applications with Yahoo! using standard technologies such as XML and RSS. Click on a link at left to learn more about our products and how we can help you.

New! - Mosey up to the Feed Trough - New RSS Page

We've updated our RSS page to include lots of basic information as well as a ginormous list of every RSS feed our API de
univer

New!

Yahoo
- this t
Webja
lets yo
web. W
them,
play th
developer network are doubly excited because webjay has an API and we like nothing better than an API. (OK: we like gummi bears better than a API. But not by much.) The Webjay API enables you to manipulate and access playlists through familiar REST-like requests. Check out the API at

Featured Applications

Maps Application Gallery
Check out cool apps built with the Yahoo! Maps API

Widget Gallery
Download more cool and beautiful Widgets for the Yahoo! Widgets runtime

Flickr Application Gallery
See the web applications and tools built using the Flickr APIs

Yahoo! Music Engine

developer.yahoo.net

All photos from Flickr!

- http://www.flickr.com/photos/altemark/39593706/
- http://www.flickr.com/photos/lorri37/88171887/
- http://www.flickr.com/photos/777/70455377/
- http://www.flickr.com/photos/nodditect/81114201/
- http://www.flickr.com/photos/dplanet/81933878/
- http://www.flickr.com/photos/merfam/715786471/
- http://www.flickr.com/photos/dankamminga/7563414/
- http://www.flickr.com/photos/angela7/765988521/

All Creative Commons (Attribution Only) Licensed

Alega Light, **Normal** & **Bold**

abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz
ABCDEFGHIJKLMNOPQRSTUVWXYZ